Parameter Estimation via the EM Algorithm for Bivariate Mixed Poisson INAR(1) Claim Count Regression Models Incorporating Correlated Random Effects

George Tzougas *[1]

[1]Department of Actuarial Mathematics and Statistics and the Maxwell Institute for Mathematical Sciences, Heriot Watt University, Edinburgh, UK

Abstract

Over the past decade, there has been a growing literature on *bivariate* (and multivariate) claim count regression models that efficiently capture the dependence between claims arising from the same policy and/or from different coverages bundled within a single policy.

In this study, we introduce a family of integer-valued autoregressive processes of order 1, denoted as BINAR(1), claim count regression models with correlated random effects for modelling the dependence structure between time series of different types of claim counts from the same and/or different coverages. For demonstration purposes, we consider bivariate mixed Poisson INAR(1) claim count regression models derived using a bivariate lognormal distribution and a Gaussian copula paired with gamma marginals as mixing densities, which we refer to as the BINAR(1)–LN and BINAR(1)–GGA claim count regression models, respectively.

Both models can be regarded as extensions of the classical bivariate Negative Binomial INAR(1) claim count regression model with a shared gamma random effect, which we refer to as the BINAR(1)–GA claim count regression model. In particular, they provide greater flexibility for modelling overdispersed bivariate time series of count data compared with the BINAR(1)–GA model. The latter is obtained by imposing the restrictive assumption of positive correlation between time series of different claim types, whereas in practice negative correlations may also arise and may therefore be of interest.

Furthermore, unlike previous copula-based count regression models, for which identifiability issues may arise when a continuous copula distribution is paired with discrete marginals, in the proposed family of models identifiability of the bivariate innovation distribution is guaranteed by imposing a unit-mean constraint on the Gamma mixing densities, which are paired with a Gaussian copula.

The main contributions of this paper are as follows:

*E-mail address: George.Tzougas@hw.ac.uk

- First, before introducing the time-series components, we present a unified framework for statistical inference via the Expectation–Maximization (EM) algorithm for the BPGA, BPLN and BPGGA regression models.
- Second, we develop novel EM-type algorithms for maximum likelihood (ML) estimation of the BINAR(1)–GA, BINAR(1)–LN and BINAR(1)–GGA regression models, which have not been explored in the literature so far. The main reason is that the joint distribution of the innovations cannot be written in closed form in these models and therefore its maximization cannot be carried out using standard numerical optimization methods.
- All the models are easily implementable and are shown to perform satisfactory when they are fitted to Local Government Property Insurance Fund data from the state of Wisconsin.

Keywords: Count data time series, Binomial-mixed Poisson INAR(1) regression models with correlated random effects, Overdispersion Â· Gaussian copula, Correlations of different signs and magnitude.